IN THE CLAIMS:

Please cancel claims 8 and 11-14 and amend the remaining claims as follows:

1-14. (Cancelled).

15. (Previously Presented) A method of clustering documents in a first dataset having first documents and a related second dataset having second documents, said method comprising:

clustering said first documents to produce first document classes;

generating a vector space model of said second documents;

classifying said vector space model of said second documents using said first document classes to produce a classified vector space model; and

determining a mean of vectors in each class in said classified vector space model to produce centroid seeds; and

clustering said second documents using said centroid seeds, such that said second dataset has a similar, based on said centroid seeds, clustering to that of said first dataset,

wherein said second dataset comprises a new, but related, based on said centroid seeds, dataset different than said first dataset,

wherein said vector space model comprises a second vector space model and said clustering of said first documents in said first data comprises:

forming a first dictionary of most common words in said first dataset; and

generating a first vector space model by counting, for each word in said first dictionary, a number of said first documents in which said word occurs,

wherein said clustering of said first documents in said first dataset is based on said first vector space model.

16. (Cancelled).

17. (Currently Amended) The method in claim 16, all the limitations of which are

incorporated herein by reference, wherein said generating of said second vector space model comprises counting, for each word in said first dictionary, a number of said second documents in which said word occurs.

18. (Currently Amended) The method in claim 17, all the limitations of which are incorporated herein by reference, further comprising:

forming a second dictionary of most common words in said second dataset;

generating a third vector space model by counting, for each word in said second dictionary, a number of said second documents in which said word occurs; and

clustering said documents in said second dataset based on said third vector space model to produce a second dataset cluster.

19. (Currently Amended) The method in claim 18, all the limitations of which are incorporated herein by reference, wherein said clustering of said second documents in said second dataset using said centroid seeds produces an adapted dataset cluster and said method further comprises:

comparing classes in said adapted dataset cluster to classes in said second dataset cluster; and

adding classes to said adapted dataset cluster based on said comparing.

20. (Previously Presented) A method of clustering documents in related datasets comprising:

forming a first dictionary of most common words in a first dataset;

generating a first vector space model by counting, for each word in said first dictionary, a number of said first documents in which said word occurs;

clustering said first documents in said first dataset based on said first vector space model to produce first document classes;

generating a second vector space model by counting, for each word in said first

dictionary, a number of said second documents in which said word occurs;

classifying said second documents in said second vector space model using said first document classes to produce a classified second vector space model;

determining a mean of vectors in each class in said classified second vector space model to produce centroid seeds; and

clustering second documents in a second dataset using said centroid seeds, such that said second dataset has a similar, based on said centroid seeds, clustering to that of said first dataset,

wherein said second dataset comprises a new, but related ,based on said centroid seeds, dataset different than said first dataset.

21. (Currently Amended) The method in claim 20, all the limitations of which are incorporated herein by reference, further comprising:

forming a second dictionary of most common words in said second dataset;

generating a third vector space model by counting, for each word, in said second dictionary, a number of said second documents in which said word occurs; and

clustering said documents in said second dataset based on said third vector space model to produce a second dataset cluster.

22. (Currently Amended) The method in claim 21, all the limitations of which are incorporated herein by reference, wherein said clustering of said second documents in said second dataset using said centroid seeds produces an adapted dataset cluster and said method further comprises:

comparing classes in said adapted dataset cluster to classes in said second dataset cluster; and

adding classes to said adapted dataset cluster based on said comparing.

23. (Previously Presented) A program device readable by machine tangibly embodying a program of instructions executable by the machine to perform a method of clustering documents

in datasets comprising:

clustering first documents in a first dataset to produce first document classes;

creating centroid seeds based on said first document classes; and

clustering second documents in a second dataset using said centroid seeds, such that said second dataset has a similar, based on said centroid seeds, clustering to that of said first dataset,

wherein said second dataset comprises a new, but related ,based on said centroid seeds, dataset different than said first dataset,

wherein said clustering of said first documents in said first dataset comprises:

forming a first dictionary of most common words in said first dataset;

generating a first vector space model by counting, for each word in said first dictionary, a number of said first documents in which said word occurs; and

clustering said first documents in said first dataset based on said first vector space model.

24-25. (Cancelled).

26. (Currently Amended) A program device readable by machine, tangibly embodying a program of instructions executable by the machine to perform said method in claim 25, all the limitations of which are incorporated herein by reference, said method further comprising generating a second vector space model by counting, for each word in said first dictionary, a number of said second documents in which said word occurs.

27. (Currently Amended) A program device readable by machine , tangibly embodying a program of instructions executable by the machine to perform said method in claim 26, all the limitations of which are incorporated herein by reference, wherein said creating of said centroid seeds comprises:

classifying said second vector space model using said first document classes to produce a classified second vector space model; and

determining a mean of vectors in each class in said classified second vector space model, wherein said mean comprises said centroid seeds.

28. (Currently Amended) A program device readable by machine, tangibly embodying a program of instructions executable by the machine to perform said method in claim 26, all the limitations of which are incorporated herein by reference, said method further comprising:

forming a second dictionary of most common words in said second dataset;

generating a third vector space model by counting, for each word in said second dictionary, a number of said second documents in which said word occurs; and

clustering said documents in said second dataset based on said third vector space model to produce a second dataset cluster.

29. (Currently Amended) A program device readable by machine, tangibly embodying a program of instructions executable by the machine to perform said method in claim 28, all the limitations of which are incorporated herein by reference, wherein said clustering of said second documents in said second dataset using said centroid seeds produces an adapted dataset cluster and said method further comprises:

comparing classes in said adapted dataset cluster to classes in said second dataset cluster; and

adding classes to said adapted dataset cluster based on said comparing.

30. (Cancelled).